

POST-DEPLOYMENT MENTAL HEALTH SCREENING INSTRUMENTS: HOW GOOD ARE THEY?

MAJ (P) Paul D. Bliese, Ph.D., Kathleen M. Wright, Ph.D., CPT Jeffrey L. Thomas*, Ph.D., Amy B. Adler, Ph.D.
United States Army Medical Research Unit-Europe
Heidelberg, Germany
CMR 442 APO AE 09042

COL Charles W. Hoge, M.D.
Division of Neurosciences
Walter Reed Army Institute of Research
503 Robert Grant Avenue
Silver Spring, MD 20910

ABSTRACT

The U.S. military has been conducting psychological screening of deploying and redeploying troops for a number of years for early identification of service members at risk for mental health problems. Currently research is being conducted to assess the validity of the clinical domains included on the screening instrument. As part of this systematic assessment, a blind validation study was conducted in March 2004 with soldiers returning from a year long deployment to Iraq. The results of this study are summarized for each clinical domain, detailing the development of a short, but valid and comprehensive screening instrument.

1. INTRODUCTION

Mental health problems are some of the most common and disabling medical conditions that affect service members. Among active duty U.S. military service members, mental disorders are the leading cause of hospitalization for men and the second leading cause for women (second only to pregnancy-related admissions). Six to ten percent of U.S. military personnel receive outpatient treatment for a mental disorder each year (Hoge et al., 2002; Garvey-Wilson et al., 2003). Over 25% of service members who receive outpatient care for mental health problems leave military service within six months, a rate that is more than two times higher than the rate of attrition following treatment for any other ICD-9 illness category (Hoge et al., 2002).

Psychiatric health is also a concern in operational and post-operational environments. During Operation Enduring Freedom (OEF) and Operation Iraqi Freedom (OIF), approximately 7% of evacuations from the theater were listed as having a primary psychiatric diagnosis. Many studies have demonstrated the strong link between

deployment experiences, especially combat, and a variety of adverse mental health, psychosocial, and occupational effects, including PTSD (15-40% lifetime rate after combat), depression, substance abuse, job loss, unemployment, divorce, and severe spouse abuse (Centers for Disease Control, 1998; Iowa Persian Gulf Study Group, 1997; Prigerson, Maciejewski & Rosenheck, 2002; McCarroll, Ursano, Liu, Thayer, Newby, Norwood & Fullerton, 2000). Available data also indicate that most service members with mental health concerns do not seek treatment, due to stigma and other barriers, although very limited research has been conducted in this area (Bray, Sanchez, Ornstein et. al. 1998; Hoge, 2003).

Given the obvious importance of mental disorders among military service members and the unusual stressors experienced during deployment, it would be desirable to have a simple and cost effective ways to identify Soldiers at risk of behavioral health problems. Toward this goal, the U.S. military has been conducting psychological screening of deploying and returning troops for a number of years. In 1996, the U.S. military mandated that psychological screening be conducted before and after deployment to the Balkans in order to facilitate early identification of mental health problems in redeploying soldiers. Although the mandate was lifted in 1999, many units participating in deployment operations opted to maintain psychological screening as a part of their before and after-care procedures. Recently, units deploying to and returning from Afghanistan and Iraq have also included psychological screening as part of their post-deployment mental health assessment in the form of the mandated DD Form 2796.

Despite the longevity of the psychological screening program, relatively little research has been conducted to determine the validity of the specific psychological screening instruments used to identify symptomatic

Report Documentation Page			Form Approved OMB No. 0704-0188		
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 00 DEC 2004		2. REPORT TYPE N/A		3. DATES COVERED -	
4. TITLE AND SUBTITLE Post-Deployment Mental Health Screening Instruments: How Good Are They?				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) United States Army Medical Research Unit-Europe Heidelberg, Germany CMR 442 APO AE 09042; Division of Neurosciences Walter Reed Army Institute of Research 503 Robert Grant Avenue Silver Spring, MD 20910				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, distribution unlimited					
13. SUPPLEMENTARY NOTES See also ADM001736, Proceedings for the Army Science Conference (24th) Held on 29 November - 2 December 2005 in Orlando, Florida.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 4	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

Soldiers. Historically, screening instruments have tended to be selected based on validity work conducted in civilian populations. For instance, one might screen for depression using an instrument developed and widely used in civilian settings. When these instruments have been applied to Soldiers, the scoring procedures developed in civilian populations are likewise applied to the Soldier populations.

It is certainly logical to develop screening instruments using established instruments from civilian research; nonetheless, independent assessments of the properties of these instruments using military populations are important for two reasons. First, it is possible that scoring criteria established in civilian populations might not be directly applicable to military settings, so validity work with military populations can address the generalizability of scoring criteria. The second reason why it is important to further assess these instruments is because scales developed in civilian populations and used in civilian settings are often lengthy and adopt complex scoring algorithms. In civilian settings, neither length nor complexity are particularly problematic, because the instruments are rarely administered to more than two or three individuals. In addition, in many civilian settings the administration and scoring can easily be computerized reducing concerns about complex scoring algorithms.

In the military, however, both length and complexity are more problematic. Instruments with 15-25 items per dimension can add up to lengthy primary screening surveys when multiple dimensions (depression, traumatic stress, alcohol, etc.) are assessed. In addition, in many screening settings, individuals (not computers) immediately score the primary screen, and given the size of the military these individuals may score thousands of surveys in a matter of days. Consequently, length and scoring complexity are more problematic because they add time and error into the screening process. Consequently, one goal of screening research is to build upon existing screening instruments and establish the validity of short (3-5 items) and easily scored scales.

2. PROCEDURE

To specifically address the issues of validity, instrument length and scale complexity, we conducted a blind-validation study within a screening assessment recommended by a senior military operational commander. The blind validation screening study was conducted in March 2004 with 1,578 Soldiers returning from a year-long deployment to Iraq. In the study, Soldiers completed a primary screen survey assessing dimensions of (a) Traumatic Stress, (b) Depression, (c) Anger, (d) Relationship problems and (e) Alcohol problems. Several of the dimensions contained items from two or more scales. For instance, the traumatic stress dimension

contained items from the DD Form 2796 (the DOD Post-Deployment Health Assessment), as well as 17 items from the PTSD Checklist (PCL) developed by the National Center for PTSD (Weathers, et al., 1993).

Using previously established criteria, Soldiers who scored positively on any one of the five clinical dimensions (a-e) were instructed to undergo a secondary screening interview with a clinical provider. Using a validated structured interview format (Sheehan et al. 1998), the clinical provider determined whether or not the Soldier required a mental health referral. To support the validation aspect of the work, however, clinical providers were kept blind to the results of the primary screen when conducting secondary interviews. In addition, a sample of approximately 20% of the Soldiers scoring negative on all domains was randomly assigned the secondary interview condition. These procedures ensured that clinical providers knew nothing about the status of the Soldier prior to conducting the interview. In this way, the primary screen could be validated by determining how well the items and scales on the primary screen corresponded with clinical providers' evaluations.

3. RESULTS

The analyses are based upon the congruence between (a) Soldiers' responses on the primary screen and (b) clinical providers' evaluations. This congruence can be summarized in terms of sensitivity and specificity. Sensitivity represents the proportion of Soldiers clinical providers identify with symptoms who were also positive on the primary screening instrument. Specificity represents the proportion of Soldiers without symptoms who had negative test results. As a point of reference, we generally consider a test to be acceptable if both sensitivity and specificity are .70 or above.

In the following sections, a summary of results is provided for each clinical dimension assessed in the screening process except alcohol problems. Alcohol problems were not assessed because Soldiers had no access to alcohol during the preceding 12 months, and alcohol questions always focus on prior behavior (drinking behavior in the last month or year).

3.1 Traumatic Stress

Results showed that the four PTSD-related items on question 12 of the DD Form 2796 did a good job of identifying symptomatic Soldiers. The stem for the question read "Have you ever had any experience that was so frightening, horrible, or upsetting that, IN THE PAST MONTH, you...(1) Had nightmares about it or thought about it when you did not want to? (2) Tried hard not to think about it or went out of your way to avoid situations

that remind you of it? (3) Were constantly on guard, watchful, or easily startled? and (4) Felt numb or detached from others, activities, or your surroundings? With two positive responses, the sensitivity of these items as a screen was 0.73, and the specificity was 0.88. Interestingly, the four items did as well as the more extensive 17-item scale (Weathers, et al., 1993) using cut-off criteria published in the scientific literature. For complete results see Bliese, Wright, Adler, Thomas & Hoge (2004).

3.2 Depression

An analyses of two depression scales (Self-Rating Depression Scale or SDS, Zung, 1965; and the Patient Health Questionnaire or PHQ, Spitzer et al., 1999) revealed that four PHQ items could be used as an effective primary screen. The stem for the items was: "Over the last 2 weeks, how often have you been bothered by any of the following problems", (1) Little interest or pleasure in doing things, (2) Feeling down, depressed, or hopeless, (3) Poor appetite or overeating, and (4) Trouble concentrating on things such as reading the newspaper or watching television. If a Soldier reported symptoms more than half the days for at least one item, the sensitivity was 0.77 and the specificity was 0.88. Further details are reported in Bliese, Wright, Adler & Thomas (2004).

3.3 Anger

Based on various anger items from the literature (e.g., Buss & Perry, 1992), three items were found to create an effective anger screen. The stem for the questions was: "During the PAST MONTH, how often have you been bothered by any of the following problems" (1) Became so angry that you have broken things, (2) Was on the verge of losing control of your anger, and (3) Flew off the handle for no good reason. When a Soldier reported Sometimes, Often or Very Often to one or more of the items, the sensitivity of the test was 0.74, and the specificity was 0.86. Further details are reported in Bliese, et al., (2004).

3.4 Relationship Problems

Two items were found to be an acceptable screen for relationship problems. The first item was "Are you having marital or relationship problems (Y or N)?" and the second item was "Our relationship is strong" (Strongly Disagree, Disagree, Neutral, Agree, and Strongly Agree). A positive response on the first item and Neutral, Disagree, or Strongly Disagree response on the second item corresponded to a sensitivity value of 0.82 and a specificity value of 0.89. Further details are reported in Bliese, et al., (2004).

4. DISCUSSION

By building a blind validation procedure into a screening program requested by an operational commander, we were able to develop four very short screens and validate the utility of an instrument in use (in the case of question 12 on the DD Form 2796 for traumatic stress). The value of this research is that it moves us significantly closer to our goal of developing a short, but valid and comprehensive screening tool that can quickly be taken by Soldiers and easily scored by technicians. Our goal is to field such an instrument on a large scale within the next year.

5. CONCLUSION

Mental health problems have been identified as the most important medical correlate of attrition from military service, as well as a concern in operational and post-operational environments. Early identification and intervention with Soldiers at risk for developing such problems will lead to referrals to the appropriate service or program and may reduce the severity and chronicity of the problem and the corresponding distress. Currently the US Army is systematically validating a psychological screening instrument that can be linked with improved prevention efforts. The recent blind validation study conducted with Soldiers returning from Iraq assessed the congruence between the scales on the screening instrument and clinical providers' evaluations. This study resulted in sensitivity and specificity analyses for each of the clinical domains on the screening instrument. Four shortened scales with acceptable sensitivity and specificity were developed and will be further validated and fielded over the next year.

6. REFERENCES

- Bliese, P. D., Wright, K. M., Adler, A. B., Thomas, J. L., & Hoge, C. W., 2004: *Screening for traumatic stress among re-deploying soldiers (U.S. Army Medical Research Unit-Europe Research Report 2004-001)*. Heidelberg, Germany: USAMRU-E.
- Bliese, P. D., Wright, K. M., Adler, A. B., & Thomas, J. L., 2004: *Validation of the 90 to 120 Day Post-Deployment Psychological Short Screen (U.S. Army Medical Research Unit-Europe Research Report 2004-002)*. Heidelberg, Germany: USAMRU-E.
- Bray R.M., Sanchez R.P., Ornstein M.L., et. al., 1998: *Department of Defense survey of health related behaviors among military personnel*. Research Triangle Institute. March 1999.
- Buss, A. H., & Perry, M., 1992: The Aggression Questionnaire. *Journal of Personality and Social Psychology*, 63, 452-459.

- Centers for Disease Control Vietnam Experience Study Group, 1988: Health status of Vietnam Veterans: Psychosocial characteristics. *JAMA*, 259, 701-7.
- Garvey-Wilson A.L., Hoge C.W., Messer S.C., Lesikar S.E., Eaton K.M., May, 2003: Diagnoses in behavioral health clinics: impact on perceived burden of mental health. In: *American Psychiatric Association. Syllabus and Proceedings from the 2003 APA Annual Meeting*.
- Hoge C.W., August, 2003: *The impact of deployment on the mental health and well-being of soldiers and their families: preliminary analysis of surveys conducted among XVIII Airborne Corps soldiers deploying to SW Asia*. 6th Annual Force Health Protection Conference, Albuquerque, New Mexico.
- Hoge, C.W., Lesikar, S.E., Guevara, R., Lange, J., Brundage, J.F., Engel, C.C. Jr., Messer, S.C., Orman, D.T., 2002: Mental disorders among U.S. Military personnel in the 1990s: association with high levels of health care utilization and early military attrition. *Am J Psychiatry*, 159: 1576-1583.
- Iowa Persian Gulf Study Group, 1997: Self-reported illness and health status among Gulf War veterans: a population-based study. *JAMA*, 277, 238-45.
- Kulka, R.A., Schlenger, W.E., Fairbank, J.A., Jordan, B.K., Hough, R.L., Marmar, C.R., & Weiss, D.S., 1991: Assessment of post traumatic stress disorder in the community: Prospects and pitfalls from recent studies of Vietnam veterans. *Psychological Assessment*, 3, 547-560.
- McCarroll, J.E., Ursano, R.J., Liu, X., Thayer, L.E., Newby J.H., Norwood, A.E., Fullerton, C.S., 2000: Deployment and the probability of spousal aggression by U.S. Army Soldiers. *Military Medicine*, 165, 41-4.
- Prigerson H.G., Maciejewski P.K., Rosenheck R.A., 2002: Population attributable fractions of psychiatric disorders and behavioral outcomes associated with combat exposures among US men. *Am J Public Health*, 92, 59-63.
- Sheehan, D.V., Lecroubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., et al., 1998: The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, 59, 22-33.
- Solomon, Z., 1989: A 3-year prospective study of post-traumatic stress disorder in Israeli combat veterans. *Journal of Traumatic Stress*, 2, 59-73.
- Spitzer, R.I., Kroenke, K., & Williams, J.B., 1999: Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. *Journal of American Medical Association*, 282, 1737-1734.
- Weathers, F.W., Litz, B.T., Herman, D.S., Huska, J.A., & Keane, T.M., 1993: *The PTSD Checklist (PCL): Reliability, validity, and diagnostic utility*. Paper presented at the annual meeting of the International Society for Traumatic Stress Studies, San Antonio.
- Zung, W. W. K., 1965: A Self-Rating Depression Scale. *Archives of General Psychiatry*, 12, 63-70.